Evolving age correlated bias in Sentiment analysis by means of embedding

Dr.S.Selvakani[1] , K. Vasumathi[2] , P.Ajithkumar[3]

[1]*Assistant Professor and Head, Department of Computer Science Thiruvalluvar University College of Arts and Science Arakkonam*
sselvakani@hotmail.com

[2]*Assistant Professor and Head, Department of Computer Applications Thiruvalluvar University College of Arts and Science Arakkonam*
kulirmail@gmail.com

[3]*PG Scholar, Department of Computer Science Thiruvalluvar University College of Arts and Science, Arakkonam, Tamilnadu, India*
PAjith2798@gmail.com

***Abstract** – Computational methodologies towards type script study are convenient popular accepting features of accessible communication, such as estimations then bias in version. Yet, contemporary educations require recognized innumerable systems of unfairness in language-based models, educating worries around the possibility of broadcasting shared preconceptions in contradiction of certain collections based on sociodemographic features (e.g., femininity, meeting, topography).*

Happening this training, we pay a efficient investigation of the submission of phonological prototypes to homework dialog on getting old. We consider the usage of growth-related languages crosswise 15 soppiness study prototypes then 10 widely-used GloVe expression embeddings besides challenge to improve unfairness complete a manner of dispensation classical working out data.

Our consequences determine that important oldness preference is fixed in the productions of several sentiment analysis algorithms then conversation embeddings. We discourse the representations' appearances in relative to production preference and in what way these representations strength be greatest integrated addicted to investigation.

Index Terms – Sentiment Analysis, Topography, Broadcasting, prototype, fragments.

I. INTRODUCTION

The idea of ageism was recognized several times ago, negative attitudes and categorizes about developing older are only now acceptance international attention. The Domain Condition Suggestion has just called for a "universal operation to struggle ageism," given the relationship between undesirable understandings almost aging and reduced condition and longevity.

Age percipience and stage preconception are focuses that require also instigated to accept responsiveness inside HCI anywhere exertion highpoints the methods that investigators and inventors incline to treat senior as a "difficult" with knowledge as a resolution, moderately than broadcasting mature as a compound and accepted measure of the lifetime. To help hostage age-related pigeon-holes everywhere knowledge use, preceding labor has underscored belongings of mature grown-ups successful accessible to dynamically.

An nearby nonetheless increasing area of awareness distresses in what way the tools and performances used to recognize online performance may promulgate public biases in contradiction of certain groups, predominantly that may remain lessened or branded. Romanticism enquiry in individual is a widespread computational style to accepting boldness, upset, and judgment in manuscript. It positions commonly used to amount opinions in produce evaluations or monetary markets, which can update and determination imprinting resolutions, dogmatic operation strategies, and mechanical economic transaction classifications.

Several computational procedures have stayed exposed to revelation communal predispositions; tools for determining romanticism disagree widely in their application, from subtracting values of element words and sayings inside a file. The incident of developmental favoritism, mechanical methods of attitude surveying immediate problems interrelated to old time of life may untruthfully report more harmful arrogances in the direction of political issues or economic moneys concerning growth-related distresses, such as Medicare and Group Confidence.

Although predisposition might branch after various foundations, we transport individual consideration to addressing bias entrenched in exercise data. Aimed at many sentimentality study utensils, the production of systems or appliance education prototypes is static essentially dependent scheduled these glossed datasets.

Computational processes are searching to not simply the dimension and quality of the fundamental datasets then also towards anthropoid collective bias that happens contained by them. We focus proceeding growth-related collective predisposition in soppiness breakdown as a incident of expending computational, algorithmic implements to study understated arrogances and attitudes. There is a developing attentiveness of age perception wide-reaching.

Age-related bias in certain devours not stayed calculated with affection to widespread sentimentality breakdown tools that are rummage-sale to mark strategic conclusions nearby produces, policymaking, savings, collective service area and service. Has nearby stopped considerable work unambiguously expected at lecturing or dropping age bias in processes.

Impending fundamental preference everywhere age has suggestions for the appropriateness of these trappings in backgrounds where boldness towards period problem, as fighting fit as the conducts that understated procedures of age percipience visible in technologies that encompass unremarkable life. The principal requests interesting the contemporary training are whether oldness bias expresses in the productivity of appliance education representations and, I beg your pardon this preconception expressions corresponding crossways universally used romanticism breakdown representations in a truthful investigation background.

Our investigation concentrations together on the conduct by romanticism investigation procedures of confrontations that stand categorical trainings of period (e.g., “timeworn” or “fledgling”) as glowing as disagreements that be located contained encodings of period (as resolute through word entrenching). Specified that background is profoundly secured to algorithmic bias, assistants require called for technologies to be calculated in the frameworks in which they function.

Addressing calculate the impression of these procedures on a text-based quantity of conferences of mature to perceive in what manner age unfairness may apparent in this true-to-life environment.

II. LITERATURE SURVEY

Shaowen Bardzell et.al [6] stated that important energies in universal computing, ICT4D, and defensible collaboration enterprise, between others, HCI is gradually engaging with troubles of social change that go beyond the speedy potentials of collaboration.. We conclude by proposing an outline of a feminist HCI methodology.

In this paper the author has given the following facilities

- Feminism seems well positioned to support HCI’s increasing awareness and accountability for its own social and cultural consequences, so developing feminist HCI is worthwhile.
- A particularly promising source is feminist social science, where researchers have practiced, theorized about, and debated scientific approaches committed to social progress.
- Whereas traditional science champions the pursuit of truth and places values out of bounds, the latter argues that socio-cultural values are inevitably bound up in scientific practice and moreover that that’s how it ought to be.

There is a reluctance to appeared alarmist in the face of the climate crisis .those who have the must education and highest professional

M.Phalguna, Dr.B.Lalitha et.al [5] started that Mobile messaging applications are growing rapidly with respect to producing Internet traffic. Such traffic when examined shows different useful insights.

In this paper the author has given the following facilities

- The usage types such as text, video and audio have different requirements in mobile computing environment. For instance more bandwidth is consumed by the applications when video is transmitted.
- This is especially important in resource intensive applications running in resource constrained networks like wireless networks. As the usage types are more and there is bulk

of information to be processed in order to estimate the resources needed for each type of service.

Create compliance and operations innovation committee (COIC) Ensure the chief data and technology officers regularity brief the COIC regarding the organization big data strategy.

TolgaBolukbasi Kai-weichang et.a1[8] started that Machine learning algorithms are optimized to model statistical properties of the training data. If the input data reflects stereotypes and biases of the broader society, then the output of the learning algorithm also captures these stereotypes. In this paper, we initiate the study of gender stereotypes in word embedding.

In this paper the author has given the following facilities

- The prejudices and stereotypes in these embedding reflect biases implicit in the data on which they were trained. The embedding of a word is typically optimized to predict co-occurring words in the corpus.
- The use of embedding in applications can amplify these biases. To illustrate this point, consider Web search where, for example, one recent project has shown that, when carefully combined with existing approaches, word vectors can significantly improve Web page relevance results.

Robert N. Butler, et.a1 [4] started that in the affluent community of Chevy Chase, recent events have revealed a complex interweaving of class, color, and age discrimination that may highlight the impact of these forces in our national life.

In this paper the author has given the following facilities

- One thing is certain: further concentration of public housing in limited sections of any city concentrating the poor or the rich or the black or the old or the young only contribute to the divisiveness of our society.
- Age, race and social class discrimination are clearly inimical to the developing human community and to the extent that our community of Chevy Chase is "closed," it is inherently disadvantaged.

Seventeen per cent of our electorate is over 65 already, but at present it is not voting as a group consequently politicians are not zealously seeking the votes of older citizens.

Shaowen Bardzell et.a1 [7] started that Feminism is a natural ally to interaction design, due to its central commitments to issues such as agency, contentment, identity, equity, empowerment, and social justice.

In this paper the author has given the following facilities

- A comprehensive introduction to issues of gender and feminism as they pertain specifically to the professional practice and theorization of interaction design.
- A generative integration of specific feminist perspectives in HCI and interaction design, that is, ways that feminism can support creative activity and novel problem solving approaches.
- Examinations of how technologies construct and perpetuate gender and the ensuing implications for the practice of design.

Conceptualization of the Feminist HCI agenda, I have referred to analogous fields, including STS, architecture, and industrial design, and I have outlined a vision of how feminism provides opportunities for the discipline.

Anthony G.Greenwald, Debbie E. McGhee et.a1 [2] started that Consider a thought experiment are shown a series of male and female faces, to which you are to respond as rapidly as possible by saying "hello" if the face is male and "goodbye" if it is female

In this paper the author has given the following facilities

- In all three experiments, target-concept stimuli for IAT measures were words or names that were associated with naturally occurring categories. This allowed possible confounding of implicit attitude differences with any other differences that existed naturally among the stimulus words or names used for the various categories.

Experiments consistently confirmed the usefulness of the IAT (implicit association test) for assessing differences in evaluative associations between pairs of semantic or social categories. The findings also suggested that the IAT may resist self-presentational forces that can mask personally.

Alec Go, Richa Bhayani et.a1 [1] started that we introduce a novel approach for automatically classifying the sentiment of Twitter messages. These messages are classified as either positive or negative with respect to a query term. This is useful for consumers.

In this paper the author has given the following facilities

- Length The maximum length of a Twitter message is 140 characters. From our training set, we calculate that the average length of a tweet is 14 words or 78 characters. This is very different from the previous sentiment classification research that focused on classifying longer bodies of work, such as movie reviews.

TABLE I : l Literature Survey

Author	Methodology	Advantages	Disadvantages
ALEGO,RICHABHAYANI	Sentiment of twitter message	Language model twitter users post message	High accuracy for classifying
ANTHONY G GREENWALD	Description are both designed	Prime-target evaluate congruence	Experiments consistently confirmed
DAVE HARELY AND GERALDINE	Complex interweaving of class	Wireless and sensor-based	Individual relationships
ROBIN BREWER	Collaboration enterprise	Workers because flexibility	Mainstream crowd work
SHAOWEN BOLUKBASI	Interaction design	A generative integration of specific	Contributions indirectly benefit

TOLGA BOLUKBASI	Model statistical propertise	Embedding reflect biases implicit	Stereotypes in both professions
ROBERT N.BUTLER	National capital housing authority	Age, race and social class discrimination	Figments of the bigoted imaginations.

That using emoticons as noisy labels for training data is an effective way to perform distant supervised learning. Machine learning algorithms (Naive Bayes, maximum entropy classification, and support vector machines) can achieve high accuracy for classifying sentiment when using this method. Although Twitter messages have unique characteristics compared to other corpora, machine learning algorithms are shown to classify tweet sentiment with similar performance.

III. PROPOSED SYSTEM:

Prospective basic preconception about period partakes suggestions for the correctness of these trappings in backgrounds someplace arrogances just before age difficulty, as healthy as the methods that restrained methods of age perception visible in equipment that infiltrate commonplace lifecycle.

As fragment of the grander conversation of algorithmic unfairness, modern effort has created to evaluate the enterprise and fundamental appliances of processes that subsidize to bias, through a request for additional experimental educations.

Several sentimentality analysis trappings are lexicon-based, which encompasses spending emotion prices of element words and languages within a article to estimate a feeling assessment for the complete. Additional collective attitude (corpus-based) is to work organizing performances consuming regarded as example script to InterCity a appliance education process. Supplementary trappings are hybrids, expending some arrangement of lexicon-based and machine-learning methods.

IV. ARCHITECTURE DIAGRAM:

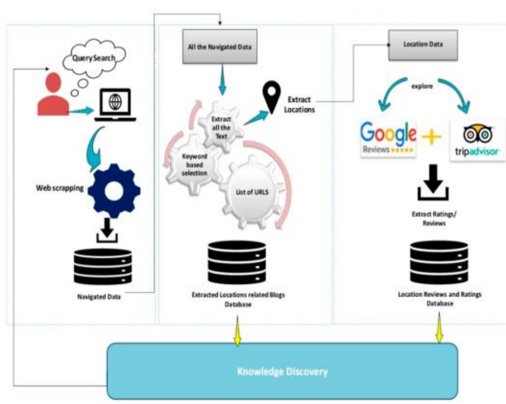

Figure1: Architecture diagram

Figure 1 is designed for the purpose of Architecture diagram with attributes such as the contain used for query search login to system and web scrapping setting to developed with navigated data is successed process of next stage all the navigated data with to the extract location all the extract all the text ,keyword based selections, list of URLS. Setting extracted locations related biogas data based used next stage knowledge discovery next step of location data two types of explore Google reviews and tripadvisor processed extract ratings/reviews used for locations reviews and ratings database successed for knowledge discovery.

V MODULES DESCRIPTION:

A. Sentiment Analysis:

Focus on age-related social bias in sentiment analysis as a case of using computational, algorithmic tools to study underrepresented attitudes and opinions. There is a growing awareness of age discrimination worldwide and age-related bias in particular has not been studied with regard to popular sentiment analysis tools that are used to make strategic decisions about products, politics, finances, social services and employment.

Sentiment analysis is one of the hottest topics and research fields in machine learning and natural language processing. The possibility of understanding the meaning, mood, context and intent of people write can offers business actionable insights into their current and future customers as well as their competitors.

B. Older Adults:

Health Information The Older Adults' Brain Scans Showed Activity In The Same Area, And Learning Life And Learning Does Not End In Old Age. Decisions By Researchers And Companies Can Be Influenced By The Relative Sentiment Of Older Adults' Experiences Compared To Younger People, Potentially Affecting The Products And Services Available To Them. Additionally, Researchers Using Sentiment Analysis to Understand Attitudes across The Lifespan Would Find those Statements Describing Older Adulthood.

The relative sentiment of older adults' experiences compared to younger people, potentially affecting the products and services available to them. Additionally, researchers using sentiment analysis to understand attitudes across the lifespan would find that statements describing older adulthood.

Changing clinical training environments into integrated geriatrics and primary care delivery systems training provides can access and addresssing age related needs of older adults and their families or caregivers at the individual, community and population levels.

Then I called the older adults based programs that will provide patients, families and caregivers with the knowledge and skills into the health age related of older adults.

C. Algorithmic Bias:

Demonstrates significant age-related bias across common sentiment analysis tools and word embedding models as well as one approach to diminishing bias in training data. The findings have implications for how researchers interpret sentiment analysis results, the strategies we use to understand and mitigate bias, and the challenges of using these techniques to study online social movements.

Algorithm bias describes systematic called and repeatable errors in a computer system that creates unfair outcomes such as privileging one arbitrary group of users over other.

Bias can emerge due to many factor that I will called the age related bias analysis relation.

D. Relation Wide Range:

The experience of aging discussions here cover a wide range of topics, such as politics, health, government, pop culture, and news, in relation to the experience of an older person. When the aforementioned sentiment analysis tools are applied to understanding the views, opinions, and experiences reported in this corpus, the sentiment output is less positive simply because the sentences describe an older person taking part in an interaction.

The relation wide range used for their content always the related bias sentiment analysis algorithm content then I small range of ageing collection the positive sentence called the age related bias sentiment analysis programing language.

VI EXPERIMENTS AND RESULTS:

PHASE:I

Substantial changes in the soppiness of unconcealed age associated keywords, entitled besides start major changes in the corn of indirectly implicit developmental keywords created finished expression embeddings**.**

GloVe conversation embeddings in the complete progression, and studied whether there stayed difference in possessions crossways the changed expression embeddings. Granting comprises the may possibly not insulate which inserting spring generated the record bias.

The presence of stage preconception popular emotion investigation prototypes the closing segment targets to validate a system to weaken that unfairness so that assistants potency motionless yield improvement of these computational styles to learning matters someplace attitudes toward time of life difficulty.

Change the preparation dataset at first recycled to construct the Gush classifier and train own convention models with this cleaned records. This consents us to deportment a additional fine-grained analysis of bias inside a particular typical and after someplace this bias activates.

This signposts that the production bias prepares definitely invent from predisposition in the brands of age-linked twitters and tin remain improved by removing these exercise.

Compare common implementation techniques that may influences bias we use 15 of the 20 sentiment analysis models implemented in sentiment that span a variety of computational techniques ,domain , and levels of complexity .we exclude the remaining five module due to a lack of variance in output scores and because one model only accepts emoticons as input. Analysis modules are often used the models are standardized to produce one of three sentiment outputs are negative (-1), neutral (0), or positive (+1).

The sentiment tools for age related bias by examining the sentiment output scores using multinomial log-linear regressions the R package nnet we build two types of multinomial log linear regressions the sentiment analysis tools in order to test for the presence of age related bias across the models in process. Analysis tools according to the types of sentiment tool used lexicon-based and corpus-based. Exponentiated coefficient values greater than one indicate that the regression models sentiment is more likely sentiment production is less and exponentiated coefficient values less than one indicate models regression sentiment.

Exclude sentences that contain the word "young" or other youth-related terms as well as complex sentences with embedded clauses or unusual grammar or sentence in a culture that deliberately hides and ignores older folks. "old age is worth waiting for" although the term old appears times across our corpus, our exclusion process results in 121 sentences from our initial sentence in each of the 121 sentences we replace the term old as well "older" and oldest" with the term "young" as well as "older" and "oldest" to provide a comparative dataset 242 sentence total. Our goal is doing this is to understand if sentiment analysis tools provide equivalent sentiment analysis younger people and youth instead of old age.

In this stage of our analysis aim to under stand whether sentences featuring keywords related to older age "old" "older", "oldest" are on average scored more negatively than the same sentences with words related to youth "young", "younger", "youngest" this difference varies depending on the particular type of model lexicon-based or corpus-based and form of validation data.

The used various sentences analysis methods age related terms in the "old" verses "young" are our independent variable of interest.

The type of sentiment analysis tools supervised learning based tools corpus based as opposed to lexicon –based were more likely to indicate either positive or negative sentiment rather than natural compared with unsupervised , lexicon-based tools , indicating a polarizing effect.

Supervised learning based tools had a polarizing effect on the likelihood of both positive and negative indications and because the sentiment analysis methods were more likely to indicate positive for "young" sentence. Analysis the data from all 15 sentiment models revealed a significant interaction between age and the types of sentiment method.

PHASE:II

Age-related bias may seep into computational approaches invarious ways .in order to better understand sources of potential bias we now turn to analyse whether age –related bias may to rooted in word embeddings encode implicit associations with age and aging. We again manipulate specific words in sentence templates, but now we generate the adjectives inserted into the templates by taking a list of common English adjectives and skewing them "old" or "young" through the use on word embedding. Word embeddings are multi-dimensional vectors (often 100-300 dimensions) where each vector represents a specific word and the values for each dimension are learned based on the context (i.e., surrounding words) within which that word commonly appears. One of the most salient emergent properties of word embeddings is that they have been shown to encode analogies (e.g., “king” – “man” + “woman” =“queen”) Thus, word embeddings can be used to transfer the relationship between two words between “man” and “woman”) onto a different word “king” and provide a reasonable semantic analog “queen”.

While word embeddings are effective at capturing semantic and syntactic properties of words, they also have been shown to latently encode stereotypes and human biases “computer programmer” – “man” + “woman” =“homemaker”) We explore this in the context of age and generate “older” and “younger” analogy of common adjectives. We start with the 500 most communal English adjectives [19] and then generate “older” and “younger” analogs for each adjective. For example, we find in one embedding that “stubborn” – “young” + “old” gives “obstinate” while “stubborn” – “old” + “young” gives

In line with the results from phase one, which found significant differences in the sentiment of explicit age related keywords, we also found significant differences in the sentiment of implicitly coded age-related keywords generated through word embeddings. The full regression results indicated that sentences constructed with implicitly “old” adjectives more negatively compared with the control adjective. Sentences with implicitly “young” adjectives were 1.09 times *more* likely to be scored positive (p<0.01, 95% CI [1.075, 1.101]). And sentences with implicitly “young” adjectives were 0.94 times as likely to be scored negatively (p<0.01, 95% CI [.926,.952]).

We included all 10 GloVe word embeddings in the full regression[6], and examined whether there was variation in effects across the different word embeddings. Although we could not isolate which embedding source yielded the most bias, the Wikipedia embeddings demonstrated the least amount of bias, whereas Twitter embeddings led to the greatest bias. When examining the individual regressions. Models indicate a significantly greater likelihood of positive sentiment in “young” adjectives as compared to the control adjective (*Adjective-Young*). In contrast, 12 of 15 models exhibit a significantly lower likelihood of indicating positive sentiment for the “old” adjectives compared to the control (*Adjective-Old*). In terms of negative sentiment, 11 of 15 models see a significantly lower likelihood of indicating negative sentiment for “young” adjectives (*Adjective-Young*).

PHASE: III

That the first two phases of our work reveal the existence of age bias in sentiment analysis models the final phase aims to demonstrate a method to diminish that bias so that researchers might still take advantage of these computational approaches to study topics where attitudes toward age matter. In this phase, we modify the training dataset originally used to create the Sentiment140 classifier and train our own custom models with this filtered data.

Each of our custom models shares the same architecture and only vary in the data that we use to train them. This allows us to directly connect output bias to changes in the train data. The architecture of each of our custom models is a Maximum Entropy bag-of-words classifier, which is a widely-used approach in various text classification problems, including sentiment analysis that predicts the most likely label (e.g. “positive” or “negative” for a given input using logistic regression. Bag-of-words models convert text inputs to a set of words, disregarding word order and grammar but retaining word frequency. This set of words is used as an input to the model, which then learns how different patterns of words map to the different labels across thousands of inputs. Create and train our models. For training data, we needed a dataset of labeled text that we could manipulate for our custom classifiers. We adopt the train data used by Sentiment140 because it is one of only two publicly-available, annotated training datasets used to train.

After randomly selecting 169 sentences containing the term “old”, we duplicate the sentences and replaced the term “old” with “young” to double the set to 338 sentences, which are then used to test the custom classifiers for the presence of bias (i.e. difference between output probabilities for “old” and “young” sentences). We increased the sample size to provide greater sensitivity and to help illuminate whether our filtering approaches could be effective. For this phase of analysis, we analyze the outputs from each of the custom-trained models using a paired t-test to determine the extent of bias that results from training on each of the different corpora.

The increase in likelihood for a sentence to be classified as “positive” when “old” is replaced with “young”. The results of each classifier output side we find the greatest output bias in classifiers trained on the *Age-Related* and *Original* corpora (both of which contain tweets with “old” and “young”) and no significant bias in the *Age-Removed* corpora. This indicates that the output bias does indeed originate from bias in the labels of age-related tweets and can be remedied by removing these training examples. The custom classifier trained on the *Original* dataset produced significant bias with respect to the terms “old” and “young” ($p<.0027$) where sentences containing the terms “old”, “older”, or “oldest” were more likely to be classified as negative. This result is in line with those of our phase one aggregated analysis. The custom classifier trained on the *Age-Related* corpus also produced significant bias ($p<.0028$). The outputs of this classifier were more negative compared to the custom classifier trained on the full Sentiment140 dataset, indicating the age-related tweets in the age related bias sentiment analysis.

VII CONCLUSION

A number of popular and diverse sentiment tools, with respect to age-related bias. The called find significant age-related bias among a variety of tools and commonly-used word embedding and successfully reduce bias in a custom-built classifier.

While the process of provide a first step in understanding how the technical characteristics of sentiment algorithms affect bias and identify one technique for reducing bias, our analysis is not exhaustive.

Future work should consider additional characteristics of algorithmic models, such as the type of classifier implemented and richer model parameters. Further, researchers should consider the unique challenges of using computational techniques such as sentiment analysis to study underrepresented assemblies and social movements.

As the "new power brokers in society," algorithms affect many aspects of life, including hiring, social policy, and finance; all of which are domains where age discrimination is common. In calculation to understanding social bias in algorithms, element can use them as a lens to understand how unrecognized social bias operates at scale.

VIII REFERENCES

[1]. Alec Go, Richa Bhayani, and Lei Huang. 2009. Twitter sentiment classification using distant supervision. CS224N Project Report, Stanford 1, 12.

[2]. Anthony G Greenwald, Debbie E Mcghee, and Jordan L K Schwartz. 1998. Measuring Individual Differences in Implicit Cognition: The Implicit Association Test. Journal of Personality and Soclal Psychology 1464–1480. https://doi.org/10.1037/00223514.74.6.1464.

[3]. Dave Harley and Geraldine Fitzpatrick. 2009. YouTube and intergenerational communication: the case of Geriatric1927. Universal Access in the Information Society 8, 1: 5–20. https://doi.org/10.1007/s10209-008-0127-y.

[4]. Robert N. Butler. 1969. Age-ism: Another form of bigotry. Gerontologist 9, 4: 243–246. https://doi.org/10.1093/geront/9.4_Part_1.243.

[5]. Robin Brewer, Meredith Ringel Morris, and Anne Marie Piper. 2016. "Why would anybody do this?": Understanding Older Adults' Motivations and Challenges in Crowd Work. Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16): 2246–2257. https://doi.org/10.1145/2858036.2858198.

[6]. Shaowen Bardzell. 2010. Feminist HCI: Taking Stock and Outlining an Agenda for Design. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10), 1301–1310.

[7]. Shaowen Bardzell and Jeffrey Bardzell. 2011. Towards a Feminist HCI Methodology: Social Science, Feminism, and HCI. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11), 675–684.

[8]. Tolga Bolukbasi, Kai-Wei Chang, James Zou, Venkatesh Saligrama, and Adam Kalai. 2016. Quantifying and Reducing Stereotypes in Word Embeddings. arXiv preprint.